

## SMBOによるレプリカ交換法のオンライン最適化

著者	土屋 遼太郎
出版者	法政大学大学院理工学・工学研究科
雑誌名	法政大学大学院紀要．理工学・工学研究科編
巻	62
ページ	1-5
発行年	2021-03-24
URL	<a href="http://doi.org/10.15002/00023978">http://doi.org/10.15002/00023978</a>

# SMBO によるレプリカ交換法のオンライン最適化

## OPTIMIZATION OF REPLICA EXCHANGE METHOD BY SMBO

土屋遼太郎

Ryotaro Tsuchiya

指導教員 平原誠

法政大学大学院理工学研究科応用情報工学専攻修士課程

In this research, the sequential model-based global optimization (SMBO) is used to sequentially adjust the parameter values of replica exchange method. In the conventional replica exchange method, the optimum number of replicas and the upper and lower limits of temperature must be determined. Therefore, herein, we attempt to reduce the parameters setting load of the user and improve the search accuracy by updating the parameter values using SMBO.

**Key Words** : *Replica exchange method, Sequential Model-Based Global Optimization, Traveling salesman problem*

### 1. はじめに

組み合わせ最適化問題には解に小さな変化を与えること（近傍操作）により改善された場合のみ、解を遷移させる単純局所探索という解法がある。しかし、これでは改悪の方向に進むことができず、局所解に陥った際に脱出することができない。そのため、改悪への遷移をある程度許しつつ短時間で良質な解を求める枠組みがあり、メタヒューリスティクスと呼ばれる。その1つとして、レプリカ交換法が提案されている。レプリカ交換法は温度の異なる複数のレプリカで同時並列的に解を探索して一定期間ごとに隣り合うレプリカ間で解を交換することで、局所解に陥りづらくし、良質な解を得ることが期待される。レプリカ交換法には最高温度や最低温度、レプリカ間の温度の比（温度比）といったパラメータを持つ。従来[1]は、最高温度と最低温度の値の決定には近傍の全探索を必要とし、問題の規模により長い処理時間を要した。また、温度比はユーザが試行錯誤して設定する必要があった。そのため、利用には障壁があると言える。

機械学習の分野においてはパラメータが多数存在し、それら値の調整が課題となっている。自動調整の手法として、Sequential Model-Based Global Optimization (SMBO) が提案されている[2]。これは確率過程回帰によりコストを回帰し、回帰結果から次に試行すべきパラメータの値を選択することで、適切な値を探索する方法である。

そこで指数分布により近傍操作によって得られる改悪の分布を近似することで最高温度と最低温度を推定する手法と SMBO によりパラメータの値を逐次調整する手法

を提案する[3][4]。性能評価には、巡回セールスマン問題 (TSP) のベンチマーク問題である TSPLIB[5]の問題を用いた。

### 2. レプリカ交換法

組み合わせ最適化問題を解く方法として単純局所探索法が存在する。これは現在の解に小さな変化を与えて候補を作り、コストが改善した場合にそれを受理する方法である。しかし、これでは局所最適解に陥った場合に抜け出すことができない。

そこで改善しなかった（改悪した）場合もある確率で受理する代表的な手法として、シュミレーテッドアニーリング(SA)がある。SAは焼きなましから着想を得た手法であり、改悪にも遷移しやすい高温の状態から、改悪に遷移しにくい低温へ徐々に冷却することで良質な解を得る手法である。よって、SAでは低温において局所最適解からの離脱が困難となる。レプリカ交換法では温度の異なる複数のレプリカで同時並列的に探索を行ない、一定周期ごとに解交換を行うことで局所最適解に陥りづらくしている。

始めに*i*番目のレプリカに注目する。コスト関数を*f*, 温度を*t<sub>i</sub>*, 現在の解を*x<sub>i</sub>*とする。このとき、近傍解*x'<sub>i</sub>*の受理確率は

$$\min\left(1, \exp\left(-\frac{f(x'_i) - f(x_i)}{t_i}\right)\right)$$

で与えられる。これにより、レプリカの温度が高いほど改悪にも移動しやすく、低いほど改悪に移動しづらくなる。次に温度の異なるレプリカ間の解交換を考える。レプリカ交換法では一定期間ごとに隣り合うレプリカ間で解交換を行う。 $i$  番目のレプリカと  $i+1$  番目のレプリカ間の解交換の受理確率は

$$\min \left( 1, \exp \left( (f(x_i) - f(x_{i+1})) \left( \frac{1}{t_i} - \frac{1}{t_{i+1}} \right) \right) \right)$$

で与えられる。この解交換により良質な解を低温のレプリカに渡してさらなる解の改善を図るとともに、局所最適解に陥った解を高温のレプリカに渡して局所最適解からの脱出を図っている。

従来研究[1]では、最低温度は最小の改悪を解交換周期内に 1 度は受理する温度とし、最高温度は最大の改悪を 50% の確率で受理する温度としていた。また、レプリカ数を 32 個とし、最低温度から等比的に各レプリカに温度を割り振っていた。

### 3. Sequential Model-Based Global Optimization

機械学習の分野では多数あるパラメータ（ハイパーパラメータ）の値の調整が課題となっている。パラメータの値を自動調整する手法の 1 つとして Sequential Model-Based Global Optimization (SMBO) がある[2]。SMBO ではベイズ最適化に基づき、あるパラメータの値でモデルを試行したことにより得られるコストを回帰し、回帰結果から次に試行すべき値を決定することを繰り返して、最適なパラメータの値を自動的に探索する手法である。

ここであるモデルのチューニングについて考える。始めに、あるパラメータの値を使用しモデルで試行した際に得られるコストを求める。次に入力をパラメータの値、教師をコストとし、Gaussian Process Regression (GPR) [6] で回帰する。この回帰結果を用い、期待する改善を表す EI 関数を最大化するパラメータの値を選択する。パラメータの値  $p$  により GPR で得られたコストの予測値を  $y(p)$ 、予測値の標準偏差を  $s(p)$ 、現在得られている最小のコストを  $\tau$  としたとき、EI 関数は

$$EI(p) = (\tau - y(p)) \Phi \left( \frac{\tau - y(p)}{s(p)} \right) + s(p) \phi \left( \frac{\tau - y(p)}{s(p)} \right)$$

で与えられる。ここで、 $\Phi$  は標準正規確率分布関数、 $\phi$  は標準正規確率密度関数である。第 1 項は GPR の予測値  $y(p)$  が小さいほど大きくなる。第 2 項は GPR の出力として得られる標準偏差  $s(p)$  が大きいほど大きくなる。これはあまり使用されておらず予測が曖昧である場合に大きくなることを表す。これにより GPR の学習に使用していない探索領域の値も採用することができる。

本研究では Latin Hyper Cube Sampling (LHS) を用い次

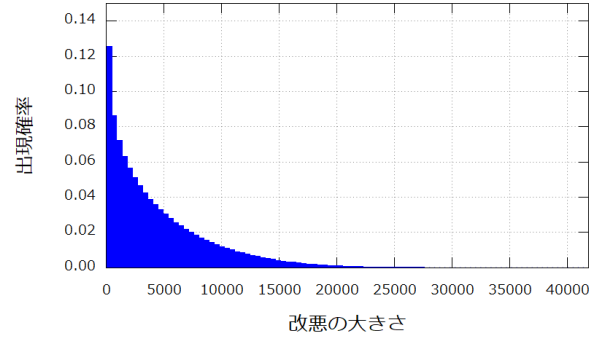


図 1. 改悪の分布

のパラメータの値の候補をサンプリングし、EI 関数を最大化する値を選択した。サンプリングしたいパラメータが 2 つある場合の LHS を考える。このとき、LHS では各次元を同じ数で昇目上に区切り、縦横被らないように区画を選択しサンプリングする。これにより、パラメータ空間内から偏りなくサンプリングすることが期待される。

### 4. 提案手法

#### (1) 最高温度と最低温度の推定

従来のレプリカ交換法[1]では最高温度と最低温度の決定のために近傍操作によって得られる改悪を全抽出し、最大の改悪と最小の改悪を求める必要があった。そこで、指数分布で改悪の分布を近似することで全抽出することなく、最高温度と最低温度を推定する手法を提案する。

始めに、TSP において有効な近傍操作といわれる 2-opt 近傍により得られる改悪を全サンプリングし、分布の確認を行う。TSPLIB の問題「rl1304.tsp」における改悪の分布を図 1 に示す。図 1 から 0 付近ほど確率が高く、指数分布に似た傾向があると考えられる。なお、解 a から解 b に遷移した場合が改悪であれば、解 b から解 a に遷移した場合は改善であるため、改善の分布は改悪の分布と y 軸に対象な分布となる。他の問題についても同様に指数分布に似た分布であることがわかった。そのため、指数分布で改悪の分布を近似することができると考えた。

指数分布の確率密度関数は

$$f(x) = \lambda \exp(-\lambda x)$$

である。パラメータ  $\lambda$  は指数分布の性質から平均値あるいは分散から求まる。本研究ではサンプリングした改悪の平均値からパラメータ  $\lambda$  を求めた。本手法では、改善と改悪ともに絶対値をとり、すべて改悪と見なす。中心極限定理から分布にかかわらず、サンプリングした平均値は正規分布に従うものと仮定する。平均値の 1 サンプルは問題の規模 100 において 30 個の平均とし、都市数  $N$  のとき 1 サンプルは

$$\frac{30}{100 \times 99 \times 98 \times 97} \times N(N-1)(N-2)(N-3)$$

個の平均とした。これにより、2-opt 近傍操作の組み合わせ総数 (母集団の規模) に比例した数をサンプリングする。そして、384 サンプルの平均値を予測結果とし、指数分布のパラメータ $\lambda$ を算出した。

次に求めた指数分布を利用し最大の改悪と最小の改悪を求めた。累積確率分布関数から、改悪が 1 つのみ現れるときを最小の改悪とし、改悪が 1 つ残してすべて現れるときを最大の改悪とした。これにより得られた最大の改悪と最小の改悪から最高温度と最低温度を従来手法に従い算出した。

## (2) SMBO によるオンライン最適化

レプリカ交換法のパラメータには主に最高温度と最低温度、レプリカ間の温度の比(温度比)がある。しかし、問題ごとに最適なパラメータの値が異なることが容易に想像される。そこで、これらパラメータの値を解の探索中に逐次調整することで局所解からの脱出が図りやすくなるのではないかと考え、SMBO によりレプリカ交換法のパラメータの値を逐次調整する手法を提案する。これにより、従来必要であった事前実験などの試行錯誤をしなくとも良質な解が得られることを期待する。SMBO では調整したいパラメータの上下限が必要であるため、4.1 節で推定した最高温度と最低温度をそれぞれ上限温度と下限温度とした。また、温度比は高温から見た比とすることで 0 より大きく 1 未満の値をとる。なお、最低温度を上昇させるとすべてのレプリカが改悪方向へ進みやすくなることや、最低温度には現時点での最良解を保持している可能性が高いことを考慮し、最低温度は SMBO による逐次最適化をしない。レプリカ数は最低温度から等比的にレプリカに温度を割り当て、最高温度を超えない最大のレプリカの数となる。

パラメータの調整方法として、機械学習のパラメータチューニングに用いられる SMBO に注目した。通常、SMBO ではパラメータの値により得られる最終的なコストをそのパラメータの値の評価とする。一方、提案手法では解交換周期ごとにパラメータを調整するため、評価を解交換周期 $k$ 回目終了時点のコスト $E_k^{(min)} = \min_{1 \leq i \leq N_k} E_{k,i}$ とする。ここで $E_{k,i}$ は解交換周期 $k$ 回目終了時点の $i$ 番目のレプリカのコストであり、 $N_k$ はレプリカ数である。なお、EI 関数の $\tau$ は解交換周期 $(k-1)$ 回目終了地点の最低コスト $E_{k-1}^{(min)}$ とする。したがって、EI 関数は解交換周期 $(k-1)$ 回目から $k$ 回目の間にどれだけ改善されるかを表す。

SMBO によるパラメータの値のオンライン最適化を加えたレプリカ交換法の手順を考える。初期状態は最高温度と温度比ともにランダムな値を初期値とし、レプリカ交換法による探索を行う。GPR への入力は更新前後の最高温度と温度比に加えて、更新前のコストの 5 つとした。更新前のコストによって改善のしやすさが変わることを考慮している。GPR の教師はパラメータの値の更新により得られるコストとした。そのため、解交換の第 1 周期終

了時点では GPR への入力と教師が揃わない。よって、第 2 周期終了時点から GPR によりコストを回帰する。GPR による回帰をしたあと、LHS でパラメータの値を複数回サンプリングし、そのうち EI 関数を最大化するものを次のパラメータの値とした。解交換の第 3 周期以降は探索の終了条件までレプリカ交換法による探索と SMBO によるパラメータのオンライン最適化を繰り返す。

## 5. 実験

### (1) 巡回セールスマン問題

組み合わせ最適化問題の 1 つに巡回セールスマン問題 (TSP) がある。TSP とは各都市を必ず 1 度だけ訪れ元の都市に戻る最短経路を求める問題である。本研究では TSP により手法の評価を行った。使用する問題は TSPLIB[3]の問題であり、都市数を $N$ 、 $i$ 番目に訪れる都市を $x_i$ 、都市 $a$ と都市 $b$ との間のユークリッド距離を $d(a, b)$ としたとき、コスト関数は

$$\sum_{i=1}^{N-1} d(x_i, x_{i+1}) + d(x_N, x_1)$$

である。なお、本実験で使用する問題は事前に大域最適解が得られている問題とした。

### (2) 指数分布による改悪の近似精度

指数分布による改悪の分布の近似精度を確認する。各問題について改悪の全サンプリングを行い、作成したヒストグラムと提案手法で近似した指数分布との距離を見る。分布間の距離を見るため Kullback-Leibler divergence (KL divergence) を使用した。2 つの確率分布 $p(x)$ 、 $q(x)$ について、 $p(x)$ の $q(x)$ に対する KL divergence は

$$D_{KL}(p \parallel q) = \int_{-\infty}^{\infty} p(x) \log \frac{p(x)}{q(x)} dx$$

で与えられる。今回は実際 (真) の各階級の出現確率と近似して得た指数分布における各階級の出現確率をみるため、KL divergence で比較する分布は離散分布である。そのため、離散分布における KL divergence を使用する。階級数が $M$ 、階級 $i$ である確率がそれぞれ $P(i)$ と $Q(i)$ であるとき、 $P(i)$ の $Q(i)$ に対する KL divergence は

$$D_{KL}(P \parallel Q) = \sum_{i=1}^M P(i) \log \frac{P(i)}{Q(i)}$$

である。KL divergence は非負であり、その値が 0 に近いほど距離が近いことを表す。加えて、推定した最大の改悪と最小の改悪の誤差についても評価した。真の値を $v_t$ 、推定した値を $v_p$ としたとき、誤差を

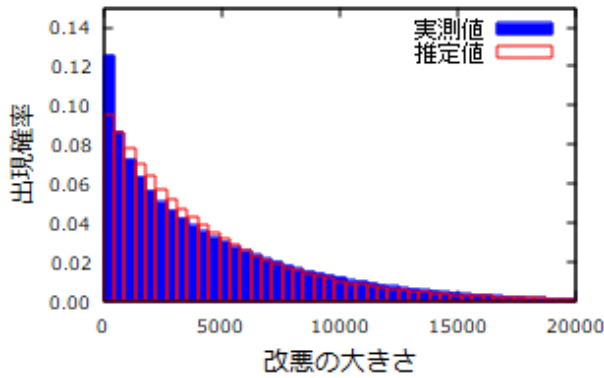


図 2. 改悪の分布と近似して得た分布

表 1. 指数分布による改悪の分布の近似精度

Problem	KL divergence	最小の改悪の誤差	最大の改悪の誤差
nrv1379.tsp	0.01686	$3.259 \times 10^3$	$5.668 \times 10^0$
pcb1173.tsp	0.01131	$7.016 \times 10^3$	$4.057 \times 10^0$
pr2392.tsp	0.01212	$2.158 \times 10^3$	$1.182 \times 10^0$
rat783.tsp	0.02684	$5.747 \times 10^4$	$1.572 \times 10^1$
rl1304.tsp	0.01077	$6.795 \times 10^3$	$-1.208 \times 10^{-1}$
rl1323.tsp	0.01268	$6.710 \times 10^3$	$-1.049 \times 10^{-1}$
u1060.tsp	0.05483	$1.225 \times 10^5$	$-1.419 \times 10^{-1}$
vm1084.tsp	0.01978	$1.663 \times 10^4$	$-2.685 \times 10^{-1}$

$$\frac{v_p - v_t}{v_t}$$

と定義した。次に、提案手法ではサンプリングして得た改悪の平均値を用いて指数分布のパラメータ $\lambda$ を求めるが、この実験では真の平均値を用いた。これにより、実際の分布が指数分布に従うか確認を行う。また、サンプリングした平均値は真の平均値を持つ正規分布に従うことが想定され、真の平均値で推定した際の評価が平均的評価であると考えられる。

問題「rl1304.tsp」における改悪の出現確率（実測値）と近似して得た指数分布における出現確率（推定値）を図 2 に示す。実際の出現確率は近似して得た分布と比べて 0 に近づくほど急峻になっており、0 付近では近似精度が悪いことが見て取れる。KL divergence は 0.0108 であった。真の最小の改悪と推定した最小の改悪はそれぞれ  $4.55 \times 10^{-13}$  と  $3.09 \times 10^{-9}$  であり、真の最大の改悪と推定した最大の改悪はそれぞれ 41522.3 と 36505.5 であった。また、最大の改悪の指定誤差と最小の推定誤差はそれぞれ 6795.2 と  $-0.12082$  であった。各問題における KL divergence、最小の改悪と最大の改悪の誤差を表 1 に示す。全体として最小の改悪の誤差が大きいのことが分かった。これは実際の分布が近似して得た分布と比べ 0 付近が急峻なためであることが原因と言える。

表 2. TSP における平均誤差

Problem	従来研究	提案手法			
		11	21	31	41
nrv1379.tsp	2.05	*3.03	*2.96	*3.12	*3.06
pcb1173.tsp	2.87	3.05	3.14	*3.28	*3.43
pr2392.tsp	3.09	3.36	*3.45	*3.87	*3.66
rat783.tsp	2.26	*2.92	*3.34	*3.19	*3.40
rl1304.tsp	5.37	*2.42	*2.92	*3.40	*3.91
rl1323.tsp	3.76	*2.36	*2.42	*3.13	*3.10
rl1889.tsp	3.73	*3.17	3.85	3.83	4.26
u1060.tsp	1.26	*2.72	*2.79	*2.99	*3.10
vm1084.tsp	3.16	3.13	3.23	3.32	3.42

### （3）TSP における提案手法の精度

提案手法と従来研究[1]との結果を比較する。従来研究と探索回数を一致させるため、解交換周期内の総探索回数は

$$N \times 20 \times 32$$

とし、終了条件は解交換周期 160 回とした。ここで  $N$  は問題の規模である。近傍操作は従来研究に従う。LHS でサンプリングする次のパラメータの候補数は 11, 21, 31, 41 の場合を試した。

各問題、各 LHS のサンプリング数について 10 回試行した際の誤差の平均値（平均誤差）を表 2 に示す。大域最適解のコストを  $E_g$ 、得られたコストを  $E$  としたとき、誤差を

$$\frac{E - E_g}{E_g} \times 100[\%]$$

とし、この平均を平均誤差とした。従来研究と有意水準 0.05 で有意差が認められたものに「\*」を付けた。「rl1304.tsp」と「rl1323.tsp」については次のパラメータの候補数をいくつにしても従来研究より平均誤差が低く、有意差が得られた。また、「rl1889.tsp」と「pcb1173.tsp」、「pr2392.tsp」、「vm1084.tsp」については次のパラメータの候補数次第では従来研究との有意差がなく、同等という結果が得られた。その他問題は従来研究と比べ平均誤差が高く、有意差が得られたため従来研究より劣るという結果になった。

### （4）真の最大・最小の改悪を用いた場合の TSP における精度

表 1 と表 2 の結果から改悪の分布の近似精度が高いほど、TSP における精度が高い傾向が見られた。そこで、真の最大の改悪と真の最小の改悪を用いて最高温度と最低温度を算出した場合の TSP における精度を確認した。これは、完全に改悪の分布を近似できた状況を意味する。ま

表 3. 真の温度における TSP の精度

Pbblem	提案手法			
	11	21	31	41
nrw1379.tsp	3.30	3.42	3.36	3.34
pcb1179.tsp	3.69	3.68	3.77	4.85
pr2392.tsp	3.24	4.59	6.00	3.70
u1060.tsp	3.48	3.40	3.65	3.60

表 4. 性質の似た問題のデータを流用した場合における TSP の精度

Pbblem	提案手法			
	11	21	31	41
rl1304.tsp	3.46	3.67	4.37	3.85
rl1323.tsp	3.24	3.36	3.34	2.71
rl1889.tsp	3.42	3.58	3.99	3.88

た, 上限温度と下限温度を除き 5.3 節の実験と同様の条件とした。

各問題について提案手法で 10 回試行した。表 3 に各問題における平均誤差を示す。5.3 節の結果 (表 2) と比較して, すべての問題において平均誤差が大きくなった。

よって, 真の最小の改悪と真の最大の改悪を用いないほうが良いことが分かった。

#### (5) 学習データの流用

ある問題で得た GPR の学習データが他の問題に流用できるか検討する。これにより, GPR の学習データの増加による SMBO のパラメータ調整能力向上を期待する。

始めに, 問題を解いた場合に得られる GPR への入力データと教師データにコストが含まれていることに注目する。コストは問題の規模 (都市数) や性質によって値が大きく異なる。そこで問題の規模や性質を打ち消す工夫として, 都市数と改悪の平均値の積でコストを割ることで正規化した。次に「rl1304.tsp」を解く場合を考える。TSPLIB では頭文字が問題のテーマを表しているため, 「rl1304.tsp」と似た問題として「rl1323」と「rl1889.tsp」が当てはまる。そこで, 提案手法でこれら 2 問をそれぞれ 5 回解き, ここで得られた GPR の学習データを「rl1304.tsp」を解く際のデータとして加えた。その他の実験条件は 5.3 節に従う。

表 4 に頭文字「rl」の問題における実験結果を示す。5.3 節の結果 (表 2) と比較して従来研究よりはよいものの, 提案手法と比べて平均誤差が大きくなり, 精度が悪化している。

## 6. 考察

表 1 と表 2 から, KL divergence が小さいほど TSP における平均誤差が低い傾向があった。そこで, 完全に近似できれば TSP における平均誤差が下がると考え, 真の最大の改悪と最小の改悪が得られた状況を仮定したが, TSP に

おいて悪化した。このことから, 改悪の近似精度を上げれば TSP における精度が向上するとは言いきれないことがわかる。推定した最小の改悪は真の値と比べて明らかに大きい従来研究の精度を上回るあるいは同等であるという結果も得られた。このことから, 従来研究通りの最低温度が低すぎると考えられる。また, 問題によっては SMBO により逐次最適化したため効率よく探索できたことが想定される。性質の似た問題を解いた際に得た GPR への入力データと教師データを流用できるか検討したが, 有効性は確認できなかった。問題間のコストの差を埋めるための正規化を試みたが, 埋めきれなかったことが原因として挙げられる。よって, 他の問題を解いて得られたデータを利用するためには, 問題の規模や性質の違いを打ち消す工夫が必要である。

## 7. 結論

指数分布による改悪の分布の近似は TSP における精度という面において一定の成果を上げることができた。加えて, SMBO によるレプリカ交換法のパラメータの値のオンライン最適化を含めた提案手法は従来以上, あるいは同等の精度を得ることができ問題が半数あった。これらのことから, 半数の問題においてパラメータの調整の手間を省き, 従来以上あるいは同等の精度を得ることができた。

## 謝辞

本研究を進めるにあたり, ご指導いただいた指導教員の平原誠准教授に心より感謝いたします。

## 参考文献

- 1) 小西健三, 屋鋪正史, 瀧和男: “温度並列シュミレーテッドアニーリング法の巡回セールスマン問題への適用と実験的解析,” 電子情報通信学会論文誌, Vol. J80-D-I, No.2, pp.127-136, 1997.
- 2) DONALD R. Jones, MATTHIAS Schonlau, WILLIAM J. Welch: “Efficient Global Optimization of Expensive Black-Box Function, Journal of Global Optimization,” Vol.13, pp.455-492, 1998.
- 3) 土屋遼太郎, 平原誠: “SMBO によるレプリカ交換法のオンライン最適化,” 第 36 回ファジイシステムシンポジウム講演論文集, pp.385-388, 2020.
- 4) 土屋遼太郎, 平原誠: “レプリカ交換法のオンライン最適化,” 第 78 回知的システム研究会, 2020.
- 5) TSPLIB: <http://elib.zib.de/pub/mp-testdata/tsp/tsplib/tsp11.html> [accessed Nov. 10, 2020].
- 6) CARL Rasmussen, CHRISTOPHER KI Williams: “Gaussian processes for machine learning,” The MIT press, 2006.